DATA, METHODS AND THEORY IN THE ORGANIZATIONAL SCIENCES

A New Synthesis

Edited by Kevin R. Murphy

First published 2022

ISBN: 978-0-367-85770-7 (hbk)
ISBN: 978-0-367-85764-6 (pbk)
ISBN: 978-1-003-01500-0 (ebk)

Chapter 13

Rebuilding Relationships between Data, Method, and Theories

Jeffrey M. Cucina and Mary Anne Nester

DOI: 10.4324/9781003015000-17

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

13

REBUILDING RELATIONSHIPS BETWEEN DATA, METHOD, AND THEORIES

How the Scientific Method Can Help

Jeffrey M. Cucina and Mary Anne Nester

When authors write manuscripts for industrial-organizational (I-O) psychology and management literatures, they are strongly encouraged to make a new theoretical contribution if they want their work to be published in the premiere outlets. This is a relatively recent trend in I-O psychology, as documented by Cucina and Moriarty (2015) who showed a dramatic increase in the use of words beginning with "theor" over the history of the *Journal of Applied Psychology* (*JAP*) and *Personnel Psychology* (*PPsych*). Editorial statements and manuscript criteria also changed over time and authors have responded by proposing lengthy theoretical models and discourse in their manuscripts. Many I-O psychologists have noted and debated the rise of "theory" in the field (Aguinis, Bradley & Broderson, 2014; Campbell & Wilmot, 2018; Cortina, 2014; Cucina et al., 2015; Kepes & McDaniel, 2013; Köhler, DeSimone & Schoen, 2020; Ryan & Ployhart, 2014). In the past, theoretical work in I-O psychology was considered the exception rather than the rule, with a greater focus on empirical data.

The emphasis on theory has a longer history within the management literature. Indeed, there is speculation that the increased role of theory in I-O psychology stems from the management literature (Aguinis et al., 2014). Journals in management, such as the *Academy of Management Journal* (*AMJ*), often require manuscripts to make a theoretical contribution and will even desk reject manuscripts that do not emphasize theory. Note that Sutton and Staw (1995) specifically singled out I-O psychology's *JAP* and *PPsych* as being at the "most empirical end of the spectrum" of management-related journals (p. 379). These journals have changed and now are more in line with management journals, such as *AMJ*, which require strong theoretical contributions to accompany empirical research.

DOI: 10.4324/9781003015000-17

At the same time that journals have increased emphasis on theory, data scientists and some I-O psychologists have begun using machine learning in applied settings, which is a methodology that largely relies on what was once known as "dustbowl empiricism." This has a resulted in a disjointed relationship between data, methods, and theory, with some I-O psychologists overemphasizing theory, others focusing on data at the expense of building scientific knowledge, and a gap between statistical advances and research practice.

There is a need in I-O psychology to rebuild the relationships between data, methods, and theory, and the aim of this chapter is to explain how these concepts should be related. We use the scientific method as a framework for this discussion. In the remainder of this chapter, we review the scientific method, describe how data, methods, and theory are related within the framework of the scientific method, and discuss two other configurations and research approaches.

Overview of the Scientific Method

Organizational researchers are no doubt familiar with the spirit and conduct of science. However, we view it as instructive to review the scientific method as it is defined and implemented in other fields of science because the recent emphasis on theory building and theoretical contributions in I-O psychology represents a significant philosophical departure from the scientific method. For instance, authors of scientific papers in I-O psychology are often expected to start a paper or research project by developing theory, often devoting significant journal space to theory exposition. There is also an expectation that the proposed theory will be supported, which results in theory-reaffirming data and results. As a result, there is less emphasis on important empirical work, studies that report interesting and practically useful results (but that lack theory development), and research in areas that do not have extensive and complex theories. In contrast, the scientific method and research practices in other disciplines often do not begin with theory but instead with an observation. The scientific method also allows for research designs that can disconfirm (as well as confirm) proposed hypotheses and theories; in fact, there is an emphasis on designing studies that would disconfirm a theory.

Many other areas of science have a formally described approach to conducting scientific research. This approach consists of an iterative seven-step process, which Cucina et al. (2014) observed to recur in their review of textbooks and coursework from various scientific fields (e.g., astronomy, physics, biology, chemistry). A depiction of the scientific method is provided in Figure 13.1. In the following sections, we describe each step in detail and explain how data, methods, and theory relate to the activities within each step.

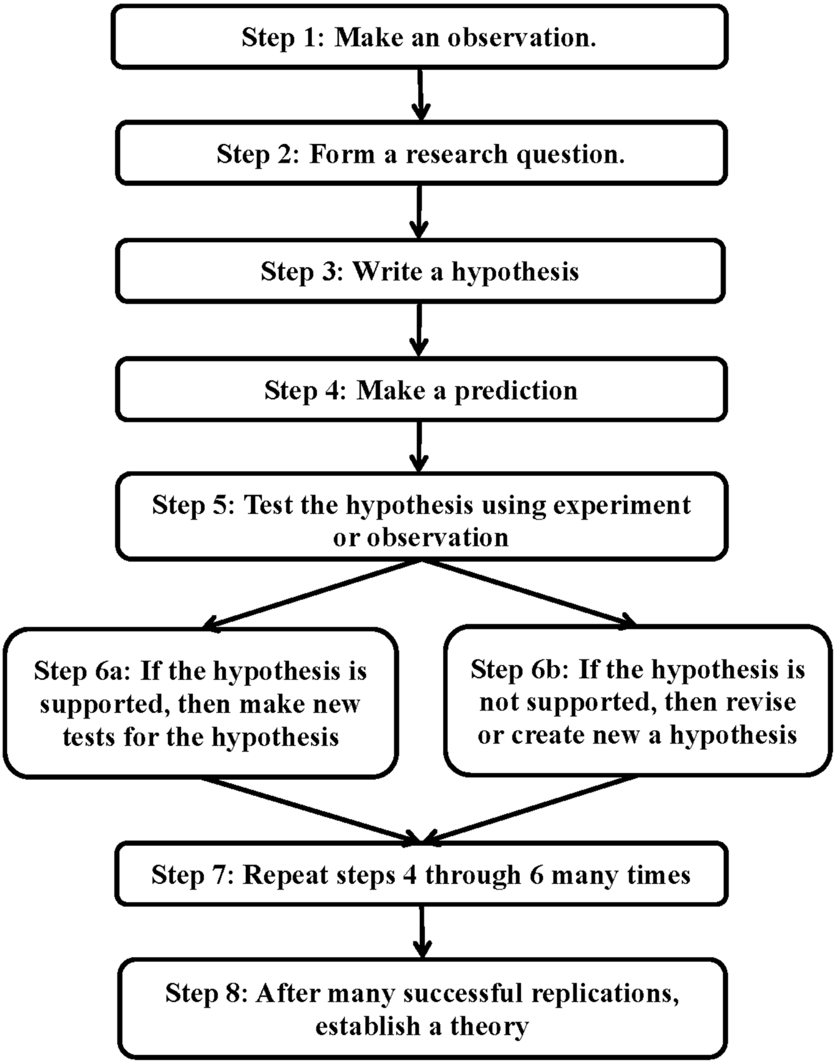


FIGURE 13.1 The scientific method as summarized by Cucina et al. (2014).

Step 1: Make an Observation

The scientific method begins with an observation about the natural world (including organizations and employees). An observation could be an anecdote, a casual observance of the everyday world, the results of previous scientific research, a case study, or information gleaned when providing consulting services to organizations. For example, a researcher could simply observe that individuals in a certain village have a lower incidence of a certain disease. Alternately, a researcher with access to a large dataset might observe a correlation between two variables. An organizational researcher might make an observation by watching the interaction between a leader and a subordinate. An observation could even be the results of a previously published study.

Although observations are most often based on data, observations based on methods and theory could serve as the basis for step 1 of the scientific method. The observations mentioned in the previous paragraph are essentially data-based observations. In some cases, the quality of the data is not robust, with small samples, uncontrolled environments, and measures with questionable construct validity limiting the conclusions that can be drawn from the study. However, this is not a problem as more robust testing of the hypothesis will occur in the subsequent steps. The observation in step 1 could also be based on more robust data (even published data and meta-analyses). For instance, a researcher might notice the correlation between two variables in a meta-analysis and build follow-up hypotheses. Observations can also be made about methods used in previous studies and a researcher could create hypotheses about using other methodologies to study an issue. Observations can also be based on existing theories, including conflicts between competing theories, or the applicability of existing theories to new areas. Observations about a stream of research can also be made. For instance, problematization (e.g., identifying incompleteness, inadequacy, and incommensurability in existing research as described by Locke & Golden-Biddle, 1997) can also serve as the basis for observations. An observation could also be based on issues facing organizations.

There are also circumstances in which a combination of data, methods, and theory can serve as the basis for an observation. For instance, Schmidt and Hunter's (2003) development of validity generalization and meta-analysis was initially built on data-based observations that showed differing validity coefficients for different industry jobs versus military data that showed consistent validity coefficients, a method-based observation that industrial studies often had much smaller sample sizes and much larger sampling error versus military studies, and a theoretical observation concerning the theory of situational specificity. This led to a hypothesis that the industrial data was inconsistent across studies due to sampling error. This was followed by the subsequent development of meta-analytic techniques, which in turn led to the theory of validity generalization and the disconfirmation of situational specificity theory.

Step 2: Form a Research Question

Most research questions in the organizational sciences focus on the relationship (e.g., correlation, regression, mean differences) between two (or more) naturally occurring variables. The research question should be concise and easily understood. Some examples include "does spatial reasoning correlate with job performance for pilots?" "are job satisfaction and turnover correlated?" and "which tasks and competencies are important for a particular job?" At this point in the process, it can be helpful to consider how meaningful and important the research question is for the field of study (e.g., the potential relevance to organizations).

Step 3: Write a Hypothesis

A hypothesis is a broad statement that aims to answer the research question. It need not be elaborately thought out; an "educated guess" (Cucina et al., 2014, pp. 358–359), conjecture, or even a hunch can serve as a hypothesis. As we will discuss later, the hypothesis could very well be revised or discarded in later steps. The hypothesis should be written so that it can apply to multiple studies, not just the study at hand. In some ways, the new theory that many I/O psychology and management journals ask authors to create in an introduction section is similar to a hypothesis. However, hypotheses are different in several regards. A hypothesis should be clear, concise, and elegant. It should not be lengthy with a large number of variables and many proposed relationships. In principle, it should be fully testable in one paper. A hypothesis could be based on existing theory, but it could also be based on an entirely new development.

At this point in the scientific method, it is beneficial to evaluate the hypothesis and potentially revise it before proceeding further. A hypothesis should make a testable proposition and it should be falsifiable (Popper, 1934, 1959). The prediction can be predicated on the logic that if the hypothesis is true, then the prediction must be true. However, Platt (1964) also makes a case for "strong inference" whereby a prediction is made such that if the prediction is true, then the hypothesis must be false. Ideally, effort should be given to the development of competing hypotheses for the research question, which would allow use of Platt's (1964) strong inference, which will be mentioned in the subsequent steps. Some elements of problematization (Locke & Golden-Biddle, 1997), especially incommensurability problematization (in which alternate theses are explored and assumptions underlying a particular area of research are examined and challenged) are also relevant here.

Step 4: Make a Prediction

A prediction is more specific in scope than a hypothesis because it is written to apply to a single study, not multiple studies. It focuses more on what is expected in the specific context of a study in terms of the methodology (e.g., participant population, organizational context, measures) rather being a general statement, such as the hypothesis. It also specifies the expected results of the data analysis. A prediction can often be portrayed formulaically (e.g., $\text{Mean}_{\text{Group A}} > \text{Mean}_{\text{Group B}}$, with a Cohen's, 1992, *d* near 1.0). The prediction is written to apply to the methodology of the study, including the measures used, the experimental manipulations, and the operational definitions of the variables.

Step 5: Test the Hypothesis Using Experiment or Observation

In this step, an empirical study is conducted to test the hypothesis. Although there are aspects of theory that relate to this step (e.g., psychometric and statistical theory are employed), data and methods play the key roles. The researcher must first design an empirical study and some thought needs to be given to the type of data that will be collected, the suitability of the data for testing the hypothesis, and the data analysis strategy. The methods used in the study are crucial in yielding suitable data for hypothesis testing. If competing hypotheses were generated in step 3, the methodology of the empirical study should yield data that could rule out one or more of the hypotheses, allowing for a "crucial experiment" (Platt, 1964, p. 347).

It can be helpful to review data from previous studies concerning the methods and data quality. If there are flaws in the research methods or the data is of poor quality, then the experiment or observational study may not adequately test the hypothesis. For psychologists, issues such as construct validity and scale reliability are critical. Many psychological measures, especially behavioral ones, are inherently unreliable. For instance, supervisory ratings of job performance have interrater reliabilities of about .52 (Shen et al., 2014; Viswesvaran, Ones & Schmidt, 1996). Many of the behavioral outcome variables used to study the relationship between personality and behavior were based on unreliable single instances of behavior. This led Mischel (1968) to question the importance of personality (vs. situational variables) in determining behavior, which had theoretical implications for personality and social psychology. However, a rebuttal by Epstein (1979) demonstrated that personality predicts behavior when more reliable behavioral outcome variables based on multiple instances of behavior are created. Thus, when designing a study, the potential validity and reliability of the measures used should be considered, lest a researcher could erroneously conclude that they have discovered a new construct or that their construct does not predict behavior.

In addition, there are other methodological considerations. For example, a power analysis should determine the sample size; however, consideration of the accuracy of effect size point estimates should also be considered. For instance, suppose a researcher anticipates a correlation of .30 between two variables. A total of 115 cases will yield 95% power for detecting the correlation; however, the 95% confidence interval about a correlation of .3 with 115 cases ranges from .12 to .46 (which is a large interval). The situation becomes worse when artifacts such as measurement error and range restriction are taken into account (Schmidt et al., 1976). Unfortunately, in our experience as reviewers of many I-O psychology manuscripts and presentations, empirical studies are too often conducted with measures having questionable construct validity, poor reliability, insufficient power[1], and insufficient precision in effect size estimates. This is unfortunate given the researcher and participant time and effort involved in conducting many empirical studies.

Once the empirical study is completed, the data are analyzed to determine if the hypothesis was supported. Obviously, quantitative and qualitative methods play an important role in this phase. Sometimes, data analysis can shed new insight on the research question and hypothesis, leading to further refinements of the hypothesis or development of *post hoc* hypotheses, as will be discussed in step 6b.

Step 6a: If the Hypothesis Is Supported, Then Make New Tests for the Hypothesis

This step is quite similar to step 5; however, researchers should attempt to test the hypothesis using additional empirical studies. Step 6a, can be implemented with multi-study articles or with separate follow-up studies (although we would recommend avoiding piecemeal publication). The new studies could be conducted by different research teams (e.g., a second research team might design a new test of a hypothesis that another team studied previously). The use of additional studies is important because Type I and II errors, unknown confounds, and experimenter error can lead to incorrect conclusions about hypotheses.

Direct replication is one approach to additional testing (Open Science Collaboration, 2015); however, modified conceptual replication often provides a better test of the hypothesis (Cucina & Hayes, 2015). In fact, modified conceptual replications are more consistent with the spirit of the scientific method since these allow for more robust testing of the hypothesis using different methodologies, types of data, measures, manipulations, populations, and so forth. For instance, directly replicating a criterion-related validation study of a scale measuring a new construct which predicted cashier job performance in a second sample of cashiers does not yield as much scientific information as a conceptual replication would. A conceptual replication might compare other approaches for measuring the construct, how well the construct adds incremental validity over other tests, whether the construct predicts performance for other jobs and for other domains of performance. Thus, in our view, parts of the recent "replication crisis" in psychology have been misguided given that an important goal for science is to develop a database of existing studies containing many different types of tests for a hypothesis. This allows for an evaluation of the external and ecological validity of the initial study's conclusion and a better determination of whether moderators exist.

Step 6b: If the Hypothesis Is Not Supported, Then Revise or Create a New Hypothesis

When a hypothesis is not supported, the best approach is to change the hypothesis, provided that the methodology and data used to test it were sufficient. This can be accomplished by revising the hypothesis or discarding it in favor of a new hypothesis. This is an instance in which data, and to some extent methods (e.g., if use of a particular methodology suggests that the hypothesis is not supported) inform a hypothesis.

In many fields, including I-O psychology, researchers are expected to predict the outcome of their studies correctly before collecting and analyzing their data. Publishing a study that does not support an initial hypothesis is discouraged and, in some cases, forbidden by the policies of academic journals (Cortina, 2016). Additionally researchers cannot revise their hypotheses (step 6b), unless they resort to hypothesizing after the results are known (HARKing, Kerr, 1998) or collecting additional data and running statistical tests multiple times to achieve significance (Simmons, Nelson & Simonsohn, 2011). There is evidence that organizational researchers sometimes alter their hypotheses after the data is collected (O'Boyle, Banks & Gonzalez-Mulé, 2017). When unreported in manuscripts, both practices are disingenuous and create statistical issues.

A much better approach is to simply allow researchers to admit their hypotheses are wrong in their papers' discussion sections and to propose (and possibly test in follow-up studies) new *post hoc* hypotheses in an open and transparent manner. Results-blind manuscript reviews and registered reports (which some I-O psychology outlets such as the *Journal of Business and Psychology* are now supporting) can serve as a foundation for this practice. Giving authors allowances (perhaps in supplemental materials) to describe additional exploratory analyses they conducted on the data can also assist in increasing transparency and providing better documentation of the research study. Additionally, the practice of maintaining a laboratory notebook[2] (Pain, 2019) often seen in the natural sciences, and open notebook science (Bradley, 2006, 2007; Schapira et al., 2019) could encourage researchers to be more transparent and provide more valuable information about their data analyses.

Many of science's greatest discoveries began with a researcher who first made a hypothesis that turned out to be flatly incorrect but later made and tested a new hypothesis that turned out to be correct. As described in Watson's (1997) firsthand account, Watson and Crick (1953) went through multiple iterations of hypotheses for the structure of DNA. They hypothesized that DNA consisted of 1, 2, 3, or 4 chains, that the chains were held together by Magnesium or Calcium ions (in fact they are held together by hydrogen bonds), that DNA's helix had rotations of different lengths (e.g., 28 or 68 Angstroms), and so forth. Their work led to so many disproven hypotheses that their laboratory director, Sir Francis

Bragg (a Nobel Laureate), attempted to ban them from continuing work on the topic. Efforts by other researchers, including Linus Pauling (Pauling & Corey, 1953), also consisted of multiple tests of disproven hypotheses. The scientific method allows for this research strategy and has led to Nobel Prize winning discoveries like Watson and Crick.

Step 7: Repeat Steps 4–6 Many Times

The seventh step in the scientific method involves repeated testing of the hypothesis in different settings, using different methodologies, and often involving multiple teams of researchers. Independent verification of findings and conceptual variations of the methodology are important practices for hypothesis testing. This helps to avoid situations in which a researcher is motivated to marshal support for a hypothesis that is credited to him or her, even if a specific study does not demonstrate support for the hypothesis.

Step 8: After Many Successful Replications, Establish a Theory

A hypothesis can only become a theory if it is supported after multiple replications (or if it can be shown to be true using logical or mathematical proofs). Whether a hypothesis rises to the level of a theory depends on the amount and quality of data from the studies that are used to test it as well as an evaluation of the adequacy of the methodology used to test it.

Establishing a theory is a group effort undertaken by an entire scientific community. It requires multiple replications in different settings to yield enough studies for a thorough meta-analysis. Even researchers who established theories without collecting their own data (e.g., Einstein's work on general relativity, meta-analysts showing the validity generalization of conscientiousness and general mental ability) make use of data from multiple studies conducted by other researchers.

Note that a single researcher would likely not follow all of the steps of the scientific method in a single study. Instead, he or she could begin with conjectures, observations, general conclusions, and other ideas and conduct and publish research on different steps of the process. For instance, a researcher could develop hypotheses or publish observations that could later be followed up and tested empirically by other research teams. Further note that abduction (i.e., proposing and testing hypotheses that explain a phenomenon) can also be implemented using the scientific method. Essentially, a researcher could use abduction to generate explanatory hypotheses which are then tested using the steps in the scientific method.

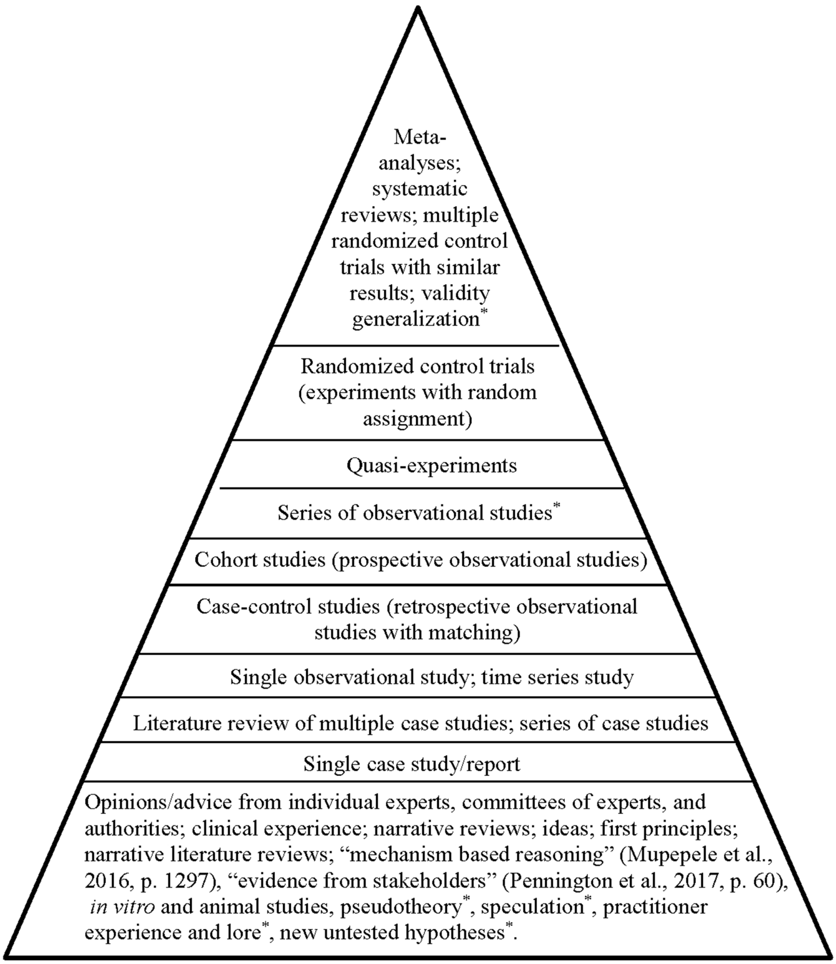
As we will discuss later in this chapter, we believe that the current state of theory development in I-O psychology and management has diverged from how theories are developed in other fields of science. I-O psychology and management researchers often start directly with creating a theory, ignoring the earlier steps in the scientific method that should lead to theory development. In addition, I-O psychology and management researchers are often encouraged to develop new theories in each paper they write as top journals often will not publish papers that do not develop new theory or that seek to test existing theories and hypotheses (Hambrick, 2007).

Designing Studies to Produce Quality Evidence and Evaluating the Quality of Evidence

When designing tests of a hypothesis, a researcher should consider the quality of evidence that will be generated by the study and how that relates to the hypothesis. Researchers should also consider the quality of evidence in support (or not in support) of a hypothesis when deciding whether to revise it or create a new hypothesis and whether the hypothesis can become a theory. As many researchers are aware, there are inherent issues with research practices and the literature, including lack of replications, fallacies (e.g., the jingle-jangle fallacy, the fallacy that correlation implies causation), and insufficient testing of hypotheses. Thus, it is important to consider the quality of evidence for a hypothesis or theory. Medical researchers have devised a pyramid or hierarchy of evidence used to evaluate evidence for medical hypotheses and theories as part of evidence-based medicine. This framework, although not without its critics (e.g., Blunt, 2015), can be used by I-O psychologists when following the steps in the scientific method. Indeed, I-O psychologists have begun to incorporate aspects of evidence-based medicine into I-O psychology and management research (Pfeffer & Sutton, 2006; Reay, Berta & Kohn, 2009).

The pyramid of evidence can be traced to a report by the Canadian Task Force on the Periodic Health Examination (1979). Since that time, different adaptations of the pyramid have been developed. Blunt (2020) has catalogued 195 versions as of August 2020, almost all of which appear in medical outlets. Reay et al. (2019) presented an adaptation of the pyramid for management research. Using many of the pyramids that Blunt (2020) catalogued, as well as the one by Reay et al. (2019), we compiled the pyramid of evidence shown in Figure 13.2. We added a few sources of evidence that are more germane to I-O psychology using asterisks.

As one progresses from the bottom of the pyramid of evidence to the top, the quality of evidence increases. At the very bottom of the pyramid are untested hypotheses and at the top of the pyramid are theories and solid empirical findings, often demonstrating evidence of causality. The pyramid shown in Figure 13.2 contains some methodologies that are not widely used in I-O psychology (e.g., case control studies) but that perhaps could be useful approaches to



Note*: Additions we made ourselves are denoted with asterisks.*

FIGURE 13.2 The pyramid of evidence as summarized from multiple sources.

studying the effects of organizational interventions when experiments are not feasible. We also should note that in some contexts (e.g., personnel selection), the ultimate goal is to establish valid prediction rather than causality and thus some forms of evidence are not always applicable.

The Role of Pseudotheorizing Within the Pyramid of Evidence

It is worth noting that the pseudotheorizing (i.e., creating an elaborate untested hypothesis; Cucina et al., 2014) that has become popular in I-O psychology and management journals is not explicitly included in the medical versions of the pyramid of evidence. It is most closely related to "mechanism-based reasoning" and "mechanistic reasoning" which Mupepele et al. (2016) included in the bottom layer of the pyramid. They define this type of reasoning as a statement that is not based on empirical data but instead is based on an inferential chain of mechanisms. Howick, Glasziou & Aronson, (2010) provide more insight on mechanistic reasoning in the medical field, where it is defined as an inferential link between mechanisms and an outcome for a patient. A mechanism is a hypothesis, theory (Howick et al., 2010), or "nomological machine" (Cartwright, 1999, p. 50) involving features or systems that have regular inputs and outputs, such as the heart. Essentially, mechanistic reasoning is conjecture that uses established findings and concepts to predict the outcome associated with an intervention. Howick et al. (2010) pointed out several examples of cases in which mechanistic reasoning led to incorrect conclusions (e.g., the famous author of childrearing advice books, Dr. Spock, used mechanistic reasoning to recommend that parents place their babies on their stomachs when sleeping to reduce the risk of babies choking on vomit and dying of sudden infant death syndrome). In our field, pseudotheorizing is not always used to justify organizational interventions; however, we do note several parallels between the quality of resulting evidence and thinking processes of mechanistic reasoning and pseudotheorizing. Both involve using the literature to make inferential leaps and create new propositions. This is not to say that mechanistic reasoning and hypothesizing should be discounted, as Howick et al. (2010) point out. It can serve as the basis for hypotheses that are later tested empirically and that ascend the pyramid of evidence. Mechanistic reasoning can also be more valid when each link in the inferential chain is tied to robust empirical evidence.

Howick et al. (2010) also noted an issue with lengthier inferential chains that has parallels with a fact that is based on path analysis. They give an example of an input having five intermediate steps, each with an effect of .90 probability leading to an outcome. On the surface, one might expect that the input will lead to the outcome with high certainty; however, the final effect of the intervention on the outcome is $.90^5$, which equals only .59. A similar situation can occur with path analysis and organizational pseudotheorizing. Suppose that an organizational intervention has a standardized path analysis coefficient of .50 with construct A, which in turn has a .50 coefficient with construct B, which in turn has a .50 coefficient with job performance and that there are no significant unmodeled paths (e.g., a direct path from A to job performance) in the model (see Figure 13.3). We might conclude that the intervention has a sizable effect on job performance because the standardized coefficients in each path are all large effect sizes (per Cohen, 1992). However, the path analysis tracing rule (Kenny, 1979, 2004) tells us that the correlation between the intervention and job performance is only $.5 \times .5 \times .5$, which equals .125, a small effect. Unfortunately, authors often

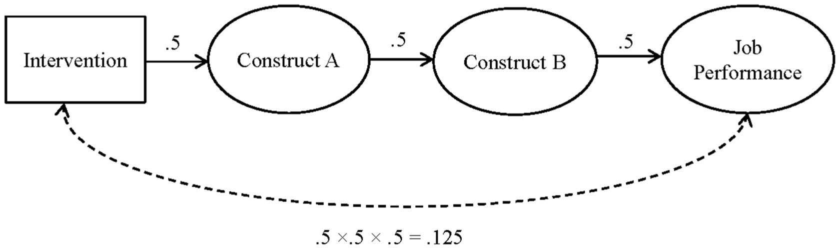


FIGURE 13.3 Path analysis model showing the diminishing effect from an intervention to job performance when it passes through two mediating constructs.

create hypotheses and pseudotheories involving these types of paths between constructs without realizing that the underlying multiplicative effect leads to small outcomes.

How the Pyramid of Evidence Informs the Scientific Method

The pyramid of evidence can inform the scientific method and there is an inherent reciprocal relationship between the two. Much of the evidence at the base of the pyramid (e.g., opinions, advice, experience, a case study) can serve as the observations in step 1 of the scientific method and can inspire the research questions and hypotheses in steps 2 and 3. Steps 4 and 5 are represented in the middle of the pyramid of evidence where single observational or experimental studies appear. As additional studies testing the hypothesis are conducted (i.e., steps 6 and 7), the evidence for a hypothesis is closer to the top. Only when the hypothesis has been supported by repeated testing can it be considered a theory.

The quality of evidence for testing a hypothesis and establishing a theory depends on the methodology used to collect data in empirical studies. Poorer quality methodologies (e.g., case studies) appear at the bottom of the pyramid of evidence. Methodologies that show a correlation between a potential cause and outcome appear in the middle, with statistical control methods (e.g., case-control studies, cohort studies) bolstering the quality of evidence. Experimental methods are required for demonstrating causality and these methods appear at the top of the pyramid. One of goals of evidence-based medicine is to identify which treatments have a causal effect on a patient's outcome, thus experiments (i.e., randomized control trials) are often considered the best type of primary study. However, issues of Type I and II errors and replicability, and other methodological issues (Cook & Campbell, 1979) can impact the results of experiments and observational studies. This is why the pyramid of evidence places repeated studies with similar results (which can include meta-analyses of those studies) at its pinnacle and the scientific method defines theory (its pinnacle finding) as a hypothesis that has received support from multiple studies using different methodologies. Thus, in most fields of science, a meta-analytic review would form the basis of theoretical establishment, whereas some I-O psychologists have written that "a review or meta-analysis does not constitute good theory" (Klein & Zedeck, 2004, p. 932).

The pyramid of evidence also adds systematic reviews, which are common in the medical literature, to its pinnacle. Uman (2011) provides an overview of medical systematic reviews, which we summarize here. Unlike a traditional narrative literature review, when conducting a systematic review, the authors search and locate articles that meet certain criteria (e.g., a specific population, specific key terms). This is similar to the literature search strategy that is used in meta-analyses. In contrast, traditional narrative reviews often do not include a detailed list of search criteria, but instead might include literature that an author is familiar with. Typically, two coders separately review each study and extract data (e.g., sample size, methodology, results) and a measure of inter-rater reliability is computed. A meta-analysis is then typically conducted with supplemental analyses (e.g., forest and funnel plots) to assist in examining heterogeneity of results and publication bias. In some cases, a meta-analysis is not appropriate due to differing methodologies or outcome variables used in the studies (Cochrane Library, n.d.). Many medical systematic reviews are conducted as part of the Cochrane Collaboration (which includes a peer review and publication of the research protocol) and often included detailed information on the methodology including separate 1–2 page tables describing characteristics of each study (e.g., methods, participants, setting, potential sources of bias, outcomes; see, e.g., Merry et al., 2012). Thus, these reviews not only provide the overall meta-analytic results but also compile summaries of individual articles in one place that could be useful to researchers and practitioners. This is one practice that I-O psychology could consider following.

Systematic reviews also consider the methodologies and limitations of the studies being reviewed, which could address some limitations of meta-analysis. For instance, if the database of studies for a meta-analysis omitted certain potential moderators, populations, or other key variables, then the authors of a systematic review could highlight those limitations for further research and to caution researchers and practitioners about potential boundary conditions for the findings. Additionally, systematic reviews are regularly updated. The Cochrane Collaboration recommends that its medical reviews be examined periodically to determine if an update is warranted and provides guidance for doing so (Cumpston & Chandler, 2020). A survival analysis of 100 completed meta-analytic systematic reviews determined that median survival time of a review (i.e., the time at which the results had a meaningful change since the previous review was conducted) was 5.5 years (Shojania et al., 2007). This speaks to the importance of (a) authors of meta-analyses examining whether their meta-analytic findings have changed since publication and (b) journals being open to publishing updated meta-analyses.

Other Configurations between Data, Methods, and Theory

The scientific method described earlier in this chapter is the predominant approach for empirical research used in nearly all fields of science (Cucina et al., 2014). I-O psychology has shifted in its approach in recent years, placing a much greater emphasis on theory (Cucina & Moriarty, 2015). However, there has also been discussion of how induction and deduction can inform scientific research. In our opinion, both approaches (i.e., focus on making a theoretical contribution and distinguishing between induction and deduction) have been misguided. In the next two sections, we discuss these two approaches in detail, pointing out how they relate to and conflict with the scientific method and how the scientific method should be the preferred approach in I-O psychology.

The Myth of the "Theoretical Contribution"

Earlier we mentioned the rise of "theory" in I-O psychology and management, which has been well documented (e.g., Aguinis et al., 2014; Cortina, 2014; Nicklin & Spector, 2016). In this section, we provide a clear definition of what theory should and should not be in science and explain how it is nearly impossible for authors to make a true theoretical contribution in their manuscripts. We make the case that by conducting research with the primary intent to make a theoretical contribution in our manuscripts; we are actually making our work less scientific. Our motivation is not to cast blame, but instead to steer I-O and management research onto the road of the scientific method and off of the road of pseudotheory.

How Do We Define Theory? The first author began with an informal review of the I-O psychology and management textbooks, editorial statements, and chapters to find a definition of "theory." It proved difficult to find a clear, concise, and agreed-upon definition of what a theory is in the I-O psychology and management literatures. Similar observations have been made by Corley and Gioia (2011) and Sutton & Staw (1995). Some of the common characteristics of what a "theory" is include long passages of text containing what Cortina (2016 p. 1143) termed "revelatory originality" rather than summaries of well-replicated empirical findings. Cortina, Aguinis, and DeShon (2017) traced the history of publications in the *Journal of Applied Psychology* and noted that in the early 1960s, introduction sections were often a couple of paragraphs. By the 1980s, a couple of pages were devoted to theory, and since the 1990s introduction sections have continued to grow in order to more fully develop it.

New "theories" are often based on case studies, common-sense, or tangentially related research. Incorporating a previously stated hypothesis or well-established theory is frowned upon and criticized as being "nothing new" and "argumentation by citation" (Ketchen, 2002, p. 586). Thus, new constructs are proposed and "surprising" statements are awarded the distinction of good theory, whereas previously supported hypotheses are not regarded as making a theoretical contribution (Corley & Gioia, 2011; Mintzberg, 2005). Figures showing complex causal models with independent, dependent, mediating, and moderating variables are common in theoretical work. This is despite the fact that the less parsimonious a causal model is the less likely it is to be true as pointed out recently by Saylors and Trafimow (2021) in the context of organizational research and by Aristotle, St. Thomas Aquinas, and William of Ockham hundreds of years ago in the context of science in general (Kaye, n.d.).

How Do Other Sciences Define Theory? Other scientific fields have much more clear and concise definitions of theory. A survey of college textbooks and interdisciplinary literature from other fields of sciences (e.g., astronomy, physics, biology) by Cucina et al. (2014) revealed a clear consensus: a theory is a well-replicated and strongly supported hypothesis. A similar survey was conducted for this paper using updated sources, and the definitions of hypothesis and theory quoted in Table 13.1 largely confirm Cucina et al.'s (2014) findings.

In I-O psychology and management, some leading journals (e.g., *Academy of Management Review*) publish only articles which essentially propose new hypotheses and models (rather than establishing that a hypothesis is well-replicated and strongly supported). Leading journals in other fields of science rarely publish articles that contain only hypotheses. For example, *Nature* (2015) publishes articles that are hypotheses "rarely, only about once a year." Moreover, it is rare for a journal to publish only speculative work. We found few analogues to the theoretical outlets of the organizational sciences in other areas of science. One exception is the journal *Medical Hypotheses* (n.d.), which publishes "interesting theoretical papers." The criteria for publication are whether or not a manuscript presents ideas that "are radical, interesting, and well-argued" (Corbyn, 2010). However, *Medical Hypotheses* has a controversial track record and is largely looked down upon by the scientific community (Corbyn, 2010).

How We Are Misdefining Theory. The theories we develop often fail to meet the definition of theory used in other sciences and many theories are only tested once, if at all (Edwards et al., 2014; Kacmar & Whitfield, 2000). Consider that theories in other fields of science are most often based on numerous empirical studies that have provided considerable support for a hypothesis. Yet much of the pseudotheory produced in the I-O psychology and management literatures lacks empirical support (Edwards et al., 2014). Modern theories in I-O psychology and management literatures often reflect the personal viewpoints of the authors rather than the very empirical evidence that should be reported. This is

TABLE 13.1 Definition of "hypothesis" and "theory" in different fields of science

Field	*Hypothesis*	*Theory*	*Source*
Astronomy	"One possible explanation of the observed facts"	"One or more 'well-tested' hypotheses [that have been] elevated to the stature of a physical law and come to form the basis of a theory of even broader applicability … Scientific theories share several important defining characteristics: they must be testable, … *tested*, … *simple*, … [and] *elegant*."	Chaisson and McMillan (2017, p. 22)
	"an idea that leads to testable predictions. The scientific method consists of observation or ideas, followed by hypothesis, followed by prediction, followed by further observation or experiments to test the prediction, and ending with a tested theory."	"If continuing tests fail to disprove a hypothesis, the scientific community will come to accept it as a theory and, after enough confirmation, eventually treat it as a law of nature. Scientific theories are only accepted as long as their predictions are borne out … Science is sometimes misunderstood because of the ways that scientists use everyday words. An example is the word *theory.* In everyday language, theory may mean a conjecture or a guess …. In everyday parlance a theory is something worthy of little serious regard. "After all," people say, "it's only a theory. " In stark contrast, a *scientific* **theory** is a carefully constructed proposition that takes into account all the relevant data and all our understanding of how the world works. ….A successful and well-corroborated theory is the pinnacle of human knowledge about the world."	Kay et al. (2013, p. 9)
Physics	"an educated guess"	"a synthesis of a large body of information that encompasses well-tested and verified hypotheses about certain aspects of the natural world."	Hewitt (2015, pp. 8–9)
	"any statement of interest that can be empirically tested. That the moon is made of cheese is a hypothesis, which was empirically tested, for example, by the Apollo astronauts."	"theories[1] are created to explain the results of experiments that were created under certain conditions …" [1]"The term 'theory' in science does not just mean 'what someone thinks,' or even 'what a lot of scientists think.' It means an interrelated set of statements that have predictive value, and that have survived a broad set of empirical tests."	Crowell (2020, p. 16)

(Continued)

TABLE 13.1 Definition of "hypothesis" and "theory" in different fields of science *(Continued)*

Field	*Hypothesis*	*Theory*	*Source*
Biology	"… [a] potential explanation… After making careful observations, scientists construct a **hypothesis**, which is suggested explanation that accounts for … observations. A hypothesis is a proposition that might be true."	"Scientists use the word **theory** in two main ways. The first meaning of theory is a proposed explanation for some natural phenomenon, often based on some general principle…Such theories often bring together concepts that were previously thought to be unrelated ….The second meaning of theory is the body of interconnected concepts, supported by scientific reasoning and experimental evidence, that explains the facts in some area of study ….To a scientist, theories are the solid ground of science, expressing ideas of which we are most certain. In contrast, to the general public, the word theory usually implies the opposite – a *lack* of knowledge, or a guess. Not surprisingly, this difference often results in confusion. In this text, theory will always be used in its scientific sense, in reference to an accepted general principle or body of knowledge …. Some critics outside of science attempt to discredit evolution by saying it is 'just a theory.' The hypothesis that evolution has occurred, however, is an accepted scientific fact – it is supported by overwhelming evidence."	Mason, Losos, Singer, Raven, and Johnson (2020, pp. 5–7)
	"a testable statement to explain a phenomenon or a set of observations"	"an explanation for a very general class of phenomena or observations that are supported by a wide body of evidence."	Freeman et al. (2017, pp. 2–3)
	"a proposed explanation for a natural phenomenon. It is a proposition based on previous observations or experimental studies ….A useful hypothesis must make **predictions** – expected outcomes that can be shown to be correct or incorrected…a useful hypothesis is testable"	"a broad explanation of some aspect of the natural world that is substantiated by a large body of evidence. Biological theories incorporate observations, hypothesis testing, and the laws of other disciplines such as chemistry and physics….In everyday language, a theory is often viewed as little more than a guess. … However, in biology, a theory is much more than a guess. A theory is an established set of ideas that explains a vast amount of data and offers valid predictions that can be tested."	Brooker et al., (2020 pp. 14–15)

(Continued)

TABLE 13.1 Definition of "hypothesis" and "theory" in different fields of science *(Continued)*

Field	*Hypothesis*	*Theory*	*Source*
	"A tentative, falsifiable explanation for one or more observations. If tests support a hypothesis, it may be incorporated into broader theories. Outside of science, *hypothesis* is used interchangeably with *theory*."	"Outside of science, the word *theory* is often used to describe an opinion or a hunch …. These tentative explanations are really untested hypotheses …. A falsifiable, comprehensive explanation for a natural phenomenon, typically backed with many lines of evidence. Nonscientists often use the term as a synonym for *opinion* or *hunch* …"	Hoefnagels (2019, pp. 13–14)
	"a proposed explanation for a phenomenon …. Commonly, when non-scientists use the word 'theory' – as in 'I've got a theory about why there's less traffic on Friday mornings than on Thursday mornings' – they actually mean that they have a hypothesis."	"an explanatory hypothesis for natural phenomena that is exceptionally well supported by the empirical data. A theory can be thought of as a hypothesis that has withstood the test of time and is unlikely to be altered by any new evidence. … it has … been repeatedly tested and no observations or experimental results have contradicted it, a theory is viewed by the scientific community with nearly the same confidence as a fact. For this reason, it is inappropriate to describe something as 'just a theory' as a way of asserting that it is not likely to be true. Theories in science also tend to be broader in scope than hypotheses … Scientific theories do not represent speculation or guesses … Rather, they are hypotheses – proposed explanations for natural phenomena – that have been so strongly and persuasively supported by empirical observation that the scientific community views them as very unlikely to be altered by new evidence."	Phelan (2015, p. 22)
Chemistry	"*a tentative explanation for a set of observations.*"	"*a unifying principle that explains a body of facts and/or those laws that are based on them* … Proving or disproving a theory can take years, even centuries, in part because the necessary technology may not be available."	Chang & Goldsby (2019, p. 5)

(Continued)

TABLE 13.1 Definition of "hypothesis" and "theory" in different fields of science *(Continued)*

Field	*Hypothesis*	*Theory*	*Source*
	"Whether derived from observation or from a 'spark of intuition,' a hypothesis is a proposal made to explain an observation. A sound hypothesis need not be correct, but it must be *testable by experiment*."	"Set of conceptual assumptions that explains data from accumulated experiments; predicts related phenomena … **theories**, based on *experiments* that test *hypotheses* about *observations* distinguishes scientific thinking from speculation …. As hypotheses are revised according to experimental results, a model emerges to explain how the phenomenon occurs …. If reproducible data support a hypothesis, a model (theory) can be developed to explain the observed phenomenon"	Silberberg & Amateis (2018, pp. 12–13)
	"a tentative and testable explanation for an observation or a series of observations"	"a general explanation of widely observed phenomena that has been extensively tested …. Theories usually start out as tentative explanations of why a set of experiments results was obtained, or why a particular phenomenon is consistently observed. Such a tentative explanation is called a **hypothesis** …. A hypothesis that withstands the tests of many experiments and accurately predicts the results of further observations and experimentation may be elevated to the rank of a scientific theory."	Gilbert et al., (2014, p. 14)
Geology	"… a tentative (untested) explanation …"	"When a hypothesis has survived extensive scrutiny and when competing hypotheses have been eliminated, a hypothesis may be elevated to the status of scientific **theory**. In everyday speech, we frequently hear people say 'that's only a theory,' implying that a theory is an educated guess or hypothesis. But to a scientist, a theory is a well-tested and widely accepted view that the scientific community agrees best explains certain observable facts."	Lutgens, Tarbuck and Tasa (2014, pp. 10)

(Continued)

TABLE 13.1 Definition of "hypothesis" and "theory" in different fields of science *(Continued)*

Field	*Hypothesis*	*Theory*	*Source*
	"A provisional explanation for observations that is subject to continual testing. If well-supported by evidence, a hypothesis may be called a theory … Tentative explanations, or hypotheses, are then formulated to explain the observed phenomenon."	"An explanation for some natural phenomenon that has a large body of supporting evidence. To be scientific, a theory must be testable (e.g., plate tectonic theory) …. Finally, if one of the hypotheses is found, after repeated tests, to explain the phenomena, then that hypothesis is proposed as a theory …. The term theory has various meanings. In colloquial usage, it means a speculative or conjectural view of something – hence, the widespread belief that scientific theories are little more than unsubstantiated wild guesses. In scientific usage, however, a theory is a coherent explanation for one or several related natural phenomena supported by a large body of objective evidence."	Wicander and Monroe (2013, p. 5)
	"tentative explanations … formulated to explain the observed phenomena."	"Finally, if one of the hypotheses is found, after repeated tests, to explain the phenomena, then the hypothesis is proposed as a theory."	Monroe and Wicander (2015, p. 6)
Psychology	"… the predicted outcome of an experiment or an educated guess about the relationship between two variables. In common terms, a hypothesis is a *testable* hunch about behavior."	"A **theory** is a system of ideas designed to interrelate concepts and facts in a way that summarizes existing data and predicts future observations."	Coon & Mitterer (2018, pp. 20–21)
	"a statement that can be used to test a prediction."	"**Theories** synthesize observations in order to explain phenomena, and they can be used to make predictions that can be tested through research. Many people believe that scientific theories are nothing more than unverified guesses or hunches, but they are mistaken (Stanovich, 2010). A theory is a well-established body of principles that often rests on a sturdy foundation of scientific evidence."	Licht and Hull (2014, pp. 20–21)

(Continued)

TABLE 13.1 Definition of "hypothesis" and "theory" in different fields of science *(Continued)*

Field	*Hypothesis*	*Theory*	*Source*
	"a tentative statement that describes the relationship between two or more *variables*. A hypothesis is often stated as a specific prediction that can be empirically tested…a testable prediction or question."	"a tentative explanation that tries to account for diverse findings on the same topic … a theory integrates and summarizes numerous research findings and observations on a particular topic."	Hockenbury, Nolan and Hockenbury (2015, p. 20)
Interdisciplinary	"A tentative explanation for an observation, phenomenon, or scientific problem that can be tested by further investigation. Scientific hypotheses must be posed in a form that allows them to be rejected."	"A plausible or scientifically acceptable, well-substantiated explanation of some aspect of the natural world; an organized system of accepted knowledge that applies in a variety of circumstances to explain a specific set of phenomena and predict the characteristics of as yet unobserved phenomena."	National Academy of Sciences (n.d.)
		"In everyday usage, 'theory' often refers to a hunch or a speculation. When people say, 'I have a theory about why that happened,' they are often drawing a conclusion based on fragmentary or inconclusive evidence. The formal scientific definition of theory is quite different from the everyday meaning of the word. It refers to a comprehensive explanation of some aspect of nature that is supported by a vast body of evidence. Many scientific theories are so well-established that no new evidence is likely to alter them substantially."	National Academy of Sciences and Institute of Medicine (2008, p. 11)
	"A tentative statement about the natural world leading to deductions that can be tested. If the deductions are verified, the hypothesis is provisionally corroborated. If the deductions are incorrect, the original hypothesis is proved false and must be abandoned or modified. Hypotheses can be used to build more complex inferences and explanations."	"In science, a well-substantiated explanation of some aspect of the natural world that can incorporate facts, laws, inferences, and tested hypotheses." Note that a Law is defined as "a descriptive generalization about how some aspect of the natural world behaves under stated circumstances."	National Center for Science Education (2016)

(Continued)

TABLE 13.1 Definition of "hypothesis" and "theory" in different fields of science *(Continued)*

Field	*Hypothesis*	*Theory*	*Source*
	"a possible (tentative) scientific explanation or prediction of an observation or set of observations…A hypothesis must be testable through a scientific investigation."	"A theory is a comprehensive explanation of some aspect of nature that is supported by a vast body of evidence. A theory is used to explain many different hypotheses about the same phenomenon or a closely related class of phenomena. Scientific theories are well-established and highly-reliable explanations that have been verified multiple times by repeated testing and have a great deal of empirical evidence that confirm them as valid."	Texas Gateway/ Texas Education Agency (n.d.)
	"an idea that proposes a tentative explanation about a phenomenon or a narrow set of phenomena observed in the natural world. The two primary features of a scientific hypothesis are falsifiability and testability, which are reflected in an 'If … then' statement summarizing the idea and in the ability to be supported or refuted through observation and experimentation."	"a broad general explanation that incorporates data from many different scientific investigations undertaken to explore hypotheses."	Rogers (2016)

Note: In all cases, emphasis appears in the original texts.

problematic as human beings have an amazing ability to create hypotheses that are flat out wrong. Examples exist in psychology (Lilienfeld et al., 2009), statistics (Lance & Vandenberg, 2015), logic (illogical fallacies such as assuming that if all S are P, then all P must be S), physics (even Albert Einstein was not infallible as he incorrectly hypothesized the existence of a cosmological constant; Harvey & Schucking, 2000; Siegel, 2013), and prescience (e.g., noting that birds did not succumb to the plague, medieval doctors assumed that wearing beak-shaped masks would make them immune to disease).

Additionally, elegance and parsimony are considered hallmarks of true scientific theories, yet much of our pseudotheories contain multiple hypotheses, paths, mediators, moderators, moderated mediators, and so forth. In some cases, prior works (e.g., previously published hypotheses or true scientific theories) are cited (and miscited) with only cursory attention to how these relate to, or can be tested in, the current study. The typical theory produced today is unclear, unconcise, untested, and often untestable. Contrast this with the true scientific theories that exist in our field such as goal-setting theory (Locke & Latham, 2002), validity generalization of general mental ability (Schmidt & Hunter, 2004), classical test theory (Nunnally & Bernstein, 1994), and item-response theory (Hambleton, Swaminathan & Rogers, 1991).

Sometimes researchers define "theory" as the "how" behind a particular process or phenomenon. In other fields of science, this is known as the mechanism of a process. Although the mechanism of how a particular medication works or selection procedure works may not be critical for practical purposes, from a basic science perspective these are questions that are important. However, simply proposing a mechanism is not an adequate approach to understanding the "how" behind a process. Empirical work would need to be conducted to test the hypothesized mechanism. Thus, the "how" behind a process is a hypothesis in and of itself.

How Do Many Management/I-O Psychology Researchers Define Theoretical Contributions? Authors today strive to make a theoretical contribution in their manuscripts. But what is a theoretical contribution? In I-O psychology and management journals, a theoretical contribution is viewed as "new and innovative ideas and insights" and "meaningfully exten[sion of] existing theory" (*Journal of Applied Psychology*, *89*(1), 178). It involves hypotheses, conjecture, stories, and imagination. Making a theoretical contribution is not viewed as an empirical process (in contrast to the scientific method). For example, consider the recommendations and inspirational messages for creating theory quoted in Table 13.2. These statements show that creating theory is much more art than science.

Unfortunately, the theoretical statements that most often appear in the recent I-O psychology and management literature are not really theories from a scientific standpoint. These statements are often poorly constructed hypotheses involving

TABLE 13.2 Quotations showing an unscientific definition of theory by OB and management theorists

Author (Year)	*Quote*	*Page*
Mintzberg (2005)	"We don't discover theory; we create it"	4
	"We get interesting theory when we let go of all this scientific correctness … and allow our minds to roam freely and creatively – to muse like mad …"	10
	"stories and anecdotes are better than measures on seven point scales and the like"	10
Davis (1971)	Good theorists are "*imaginative*" (italics appear in original)	344
	"Qualitative correlations" form better theories and are more "interesting than quantitative correlations"	323
	Too much focus on the scientific method results in "the 'Mediocre'"	328
	"'the creative spark'"	328
Sutton & Staw (1995)	Those who are "good at theoretical" work are "dreamy"	380
Klein & Zedeck (2004)	"Assertion and even evidence are no substitute for explanation and interpretation"	932
Shepherd & Suddaby (2017)	"compelling theories are at their core compelling stories." "theory building [can be centered] around the five key elements that inform every great story: conflict, character, setting, sequence, and plot and arc."	60 60

models that are too complex[3] to test in a single study, regardless of how well designed the study may be. Future researchers are discouraged from testing these hypotheses further because in doing so they fail to make their own new "theoretical contribution" by proposing new hypotheses. Thus, hypotheses are often only tested partially in the paper that proposes them (and some are not tested at all). As a result, we are creating new hypotheses and not allowing them to become falsifiable (one of the hallmarks of well-constructed hypotheses in science). These hypotheses often take up considerable journal page space, which is a limited resource. Highhouse (2014) has noted a substantial increase in the length of articles in I-O journals. This raises a question about the value of journal pages that present ideas lacking evidence of support or disconfirmation. If the goal of I-O psychology and management research is to improve the functioning of the workplace, shouldn't our journals strive to focus on publishing text that has an empirical basis? A better approach, as will be described below, would be to remove much of the hypothesizing and theoretical conjecture in favor of more concise and factual articles. The remaining pages could then be replaced with additional articles, especially conceptual replications of earlier works (Kepes & McDaniel, 2013).

What Is a True "Theoretical Contribution"? I-O psychologists' and management researchers' definition of a theoretical contribution differs greatly from that used in other sciences. In other fields, a theoretical contribution would include the results of a multi-study empirical research program that follows the scientific method. Locke & Latham's (2002) goal-setting theory is an excellent example of a theoretical contribution, yet making this contribution involved a "35-year odyssey" (p. 705) with over 400 studies by many scientists (Locke, 2007). Other examples of a theoretical contribution would include a meta-analysis showing a robust effect size (with small credibility intervals and no publication bias) or a review that convincingly shows support for a hypothesis. Occasionally, a crucial study that disconfirms an existing theory arises. This also could lead to a theoretical contribution; however, it might need to be replicated before rising to the level of a theoretical contribution. Thus, according to the scientific definition of theory and the scientific method, it would be extremely difficult for an author to make a theoretical contribution in a single paper.

In fact, in other fields of science, the term "theoretical contribution" rarely appears in individual articles. A search of two premiere multidisciplinary scientific journals (i.e., *Nature* and *Science*) was conducted for the phrase "theoretical contribution" through 2020. In *Nature*, this phrase appeared in only 101 of 422,374 articles. The term was almost always used to commend the work of someone else. Most of the hits, (i.e., 83) bestowed the term "theoretical contribution" on someone else, often in obituaries and stories of individuals receiving awards. Nine articles stated that based on existing theory, something can make a "theoretical" contribution to a process (e.g., there might be a "theoretical contribution of the three amino-acids" [Burgus et al., 1970] in a particular reaction). Seven articles mentioned the phrase in announcements for other (often new) journals. One article complained about untested theoretical contributions and the last mentioned it in terms of determining authorship credit.

The term was less commonly used in *Science*. Out of 297,556 articles, only 13 included the term. Again, the phrase "theoretical contribution" was most often used (in 11 of the 13 hits) to commend work of someone else, especially in obituaries and announcements of awards that have been given to certain individuals. Only two articles actually stated that the authors made a theoretical contribution. In an article about the planet Jupiter, Trafton and Wildey (1970) wrote that "The primary theoretical contribution of the study reported here has been the incorporation of the ammonia bands at 10 and 16 A into the model…." More recently, Lacour and Green (2014) authored an article on transmission of gay equality and stated "Our theoretical contribution is to introduce the distinction between active and passive contact, which are posited to produce different…" As an aside, this article was later retracted from *Science*.

Why Making a True Theoretical Contribution in One Paper Is Almost Mythical Authors are encouraged to write papers that make a theoretical contribution. As described above, this involves proposing a theory that is new and surprising, one that has not been presented in the literature previously. However, from a scientific standpoint, a paper that truly makes a theoretical contribution is almost mythical. Science is an incremental process, and few papers really can establish a brand new scientific theory that other scientists did not anticipate. Scientific theories evolve over many years and involve multiple published and presented studies, followed by a meta-analysis or review. In contrast with other fields of science, our field does not view this type of evidence as theory (Klein & Zedeck, 2004).

Clearly, we need to (re)embrace meta-analytic methods and other forms of credible evidence (e.g., systematic reviews) as a path toward establishing a theory. However, we also need to reconsider the notion that a theoretical contribution must be something completely new. This expectation has led to a proliferation of untested theories in I-O psychology and management (Edwards et al., 2004). Given the multiple steps involved in the scientific method for replication and designing new studies to test a hypothesis, it is often not logistically feasible for a researcher to create a novel hypothesis and establish it as a theory in a single paper, especially a primary study. Indeed, it is quite rare for a researcher to create a hypothesis and then find multiple existing studies showing enough support to establish it as a theory. Some possible exceptions exist. For example, Schmidt & Hunter's (1977) validity generalization of general mental ability tests and Einstein's theory of special relativity are possible exceptions. Yet, these theoretical contributions, although later shown to be strongly supported, underwent extensive empirical testing after being introduced. Schmidt and Hunter encountered substantial resistance to validity generalization, but after many follow-up articles (e.g., Schmidt et al., 1985) it was eventually accepted by most of the scientific community (Schmidt, 2015). Einstein's work also received initial skepticism and follow-up testing (Brush, 1999; Goldberg, 1987), but now is widely accepted. Both theories were also based on existing experimental evidence. Validity generalization was initially based on observations that small sample industrial validity coefficients varied widely, whereas large sample validity coefficients from military studies were more consistent (Schmidt, 2015).

It also might be possible for a researcher to spend years iterating through the steps of the scientific method to turn his or her new hypothesis into a theory and then publish all of the results at once. However, this is not a wise career move; the researcher's *CV* would have a huge hole, funding institutions and employers would wonder what progress is being made, and so forth. Additionally, other researchers would not be able to conduct independent tests and provide evidence that the results replicate in other contexts. Thus, this is not a viable approach.

The Best That Can Be Done. So what is the best that can be done in a single paper with respect to theory? A researcher can create a hypothesis that is new to the literature and either test it or propose that others test it (e.g., in an outlet such as *AMR*). After extensive testing this hypothesis could become recognized as a theory and the researcher could be noted for making a theoretical contribution. Additionally, a researcher could revise an existing hypothesis (and optionally test it). Researchers can also continue testing previously published hypotheses, which may one day become theory. Hypotheses can become theories when there is a large body of empirical evidence providing their support. Meta-analyses and literature reviews summarize the results of multiple studies testing a hypothesis. These types of reviews (quantitative and qualitative) provide the strongest basis for establishing a new theory, although many in our field would incorrectly view them as not making a theoretical contribution. Essentially, making a theoretical contribution, as defined in the sciences, is most often an effort of the scientific community rather than an effort of one person or one paper.

How to Use Theory Correctly. Currently, many authors develop a new theory to generate hypotheses for I-O psychology and management empirical research. However, there is another way to incorporate theory when designing research studies. There are many well-supported truly scientific theories in the basic research literature. Our field is sitting next to a scientific theory goldmine in other psychology journals, yet we neglect it in favor of creating our own untested theories. Originally, the goal of applied psychology was to apply theories and findings from basic psychology to the real world. However, we have drifted away from this goal.

Some examples of how to incorporate theory from other disciplines exist in the I-O related literature. For example, there is strong evidence for the theory of the self-serving bias in the social psychology and personality literatures (e.g., Dunning, Perie & Story, 1991). The essence of this theory is that when asked which traits are important for success, individuals have a tendency to report that the traits they perceive themselves as having are more important than other traits. Three studies have tested the theory of the self-serving bias in a new context, job analysis ratings, and have found that it exists in job analysis data (Aguinis, Mazurkiewicz & Heggestad, 2009; Cucina et al., 2012; Cucina, Vasilopoulos & Sehgal, 2005). In another example, based on empirical work encompassing over 460 datasets, Carroll (1993) developed the Three-Stratum Theory of mental abilities. One of these abilities, Meaningful Memory, bares a strong resemblance to the process of learning material in a training environment. It involves the ability to recall, after a study period, material that involves meaningful interrelations (e.g., a story, a concept, a biography of an individual). Recent work applying this ability from Carroll's theory shows that it is one of the few specific abilities that uniquely predicts training performance beyond general mental ability (Cucina et al., 2014). A third example resides in work by Carter et al. (2014) who applied the item response theory generalized graded unfolding model (GGUM; Roberts, Donoghue & Laughlin, 2000) from the psychometrics literature to scoring personality tests.

The Scientific Method as an Alternative to "Inductive" and "Deductive" Research

In recent years, there have been calls for renewed interest in inductive research within I/O psychology and OB (Locke, 2007; Locke, Williams & Masuda, 2015; Ones et al. 2017; Spector et al, 2014; Woo, O'Boyle & Spector, 2017). This can be contrasted with the heavy focus on deduction and theory building in many academic journals (Aguinis et al., 2014; Campbell & Wilmot, 2018; Cortina, 2014; Dilchert, 2017; Hambrick, 2007; Highhouse, 2014; Kepes & McDaniel, 2013; Olenick et al., 2018; Ones et al., 2017; Schneider, 2018). In this context, induction is often viewed as making an observation that is based on empirical data and then developing a general conclusion or a theory. Deduction is often viewed as developing general conclusions or theory and testing those empirically. In brief, induction is viewed as going from the particular to the general and deduction is viewed as going from the general to the particular.

In this section, we briefly review existing conceptualizations of inductive and deductive research often used in I/O psychology and OB. We then present alternative conceptualizations of induction and deduction that are more closely aligned with those appearing in the philosophical and logical literatures. We then describe how the different conceptualizations fit into the scientific method. We go further by suggesting that researchers can avoid having to make distinctions between induction and deduction by simply following the scientific method. Aspects of both conceptualizations of induction and deduction are folded into and encompassed in the scientific method. Later, we explain how our work relates to the recent debate concerning the role of and emphasis on theory in I/O psychology and OB research. (See the four articles in the point/counterpoint section of the *Journal of Organizational Behavior* edited by Nicklin and Spector, 2016, for an example of the debate). We conclude with recommendations for researchers.

Deductive and Inductive Reasoning

The roots of inductive and deductive reasoning lie in formal logic and the philosophy of science. Both induction and deduction involve using premises (e.g., statements, principles, evidence, observations) to reach a conclusion. In deductive reasoning, provided that the premises are true, a valid conclusion is necessarily true with 100% certainty. For example, suppose that we are given two premises: "All nonprofit organizations are groups organized for a purpose other than generating a profit" and "G Group is a nonprofit organization." Assuming the premises are true, we can validly conclude (with complete certainty) that G Group was organized for a purpose other than generating a profit. Logicians define induction as using evidence to reach a conclusion that may be true but that is not guaranteed to be true (Hawthorne, 2017; Skyrms, 2000). The conclusions in inductive reasoning are valid but not completely certain; instead there is a level of probability or confirmation associated with each conclusion. For example, suppose that we are given two premises: "67% of nonprofit organizations are public charities" and "G Group is a nonprofit organization." Based on the premises, we can validly conclude that there is a .67 probability that G Group is a public charity; thus, although it is likely that G Group is a public charity, we are left uncertain.

In their comprehensive discussion of the relationship between deduction and induction, Colberg, Nester and Trattner (1985) showed that these two types of reasoning converge in terms of their forms but differ with respect to the certainty of their conclusions. The examples from the preceding paragraph are presented below to show both the convergence of the reasoning forms and the difference in the certainty of the conclusion. (They are presented in conditional form rather than in set form.)

If an entity is a nonprofit organization, then it was organized for purposes other than generating a profit.

G Group is a nonprofit organization.

Therefore, G Group was (necessarily) organized for purposes other than generating a profit. (deductive)

If an entity is a nonprofit organization, then there is a .67 probability that it is a public charity.

G Group is a nonprofit organization.

Therefore (with a probability of .67), G Group is a public charity. (inductive)

In the first example above, G Group must have been organized for purposes other than generating a profit because of its membership in the set of nonprofit organizations, all of which were organized for purposes other than generating a profit. This example represents a basic form in deductive logic, called "modus ponens," in which a necessary conclusion is drawn about an individual based on information in a universal premise. It begins with the premise "if p then q." If it is affirmed that p is true, then it follows that q is true. In this example, a necessary conclusion is drawn about an individual based on information in a universal premise. In the second example, there is no universal premise. Instead, the first premise says that if an entity is a nonprofit organization, then there is a .67 probability that it is a public charity.

The valid conclusion is stated with a probability of .67, which means that it is possible but not certain that G Group is a public charity. As Colberg et al. (1985) explain:

> When logicians state that an inductive conclusion is never certain, they mean that there is no identity between the premises and the conclusion. In a deductive conclusion there is such an identity. This is the same as saying that in a deductive conclusion, if the premises are true and the schema

> is correctly constructed, then the conclusion cannot be false, whereas in an inductive conclusion, even if the premises are true and the schema is correctly constructed, the conclusion can still be false. The falsity of the conclusion in an inductive schema is compatible with the premises and with the schema. Thus, a deductive conclusion is always necessary, never probabilistic, and an inductive conclusion is always probabilistic, never necessary. (p. 682)

As a consequence of the identity between premises and conclusion in deductive arguments, they "… are usually limited to inferences that follow from definitions, mathematics, and rules of formal logic" (*Internet Encyclopedia of Philosophy*, 2016). By contrast, scientific research tries to go beyond deductive argument by observing natural phenomena. When these observations are about sets that cannot be known in their entirety, observations are usually made about a sample from the phenomenon being investigated. Any general conclusion drawn from such samples is an inductive, probabilistic conclusion. The statistical tools that enable scientists to estimate, for example, a population value from sample statistics have been developed mathematically, that is to say deductively. The use of these estimates to describe the population in question, however, represents an inductive exercise because the estimates are based on incomplete information about the population.

In the organizational science and I/O psychology literatures, deduction is often defined as applying a general principle or rule to a specific case, and induction is often described as observing a series of specific cases and inferring and formulating a general principle from the specific cases. Colberg et al. (1985) traced this misconception of inductive reasoning to early psychometric research by Thurstone (1938), who defined the term as "find[ing] a rule or principle for each item in the test."[4] It has also made its way into the research methods literature, whereby the process of first observing data and then making a general statement (e.g., a hypothesis or theory) is defined as inductive research and the process of first making a general statement (especially a hypothesis) and then observing data is defined as deductive research. This way of defining inductive and deductive research is incompatible with definitions used in logic and the philosophy of science, and it obscures the fact that all scientific research involves both induction and deduction.

The Scientific Method Incorporates the Process of Going From Particular to General and Vice Versa

The previously described distinction between inductive and deductive research in terms of general to particular and particular to general research can be rendered moot by adoption of the scientific method, which incorporates both processes.

The process of going from particular to general appears in several steps in the scientific method. A researcher could observe a particular phenomenon in step 1 (i.e., make an observation) and then construct a research question and hypothesis in steps 2 and 3 which later lead to a general conclusion in step 8 (i.e., form a theory). For instance, a researcher might begin with the observation that scores on certain measures in a battery she administered are related to leadership success. The finding could be used to form more general research questions and hypotheses concerning the constructs measured in the battery. The working hypothesis could then be tested in a new study. After repeated replication and meta-analysis, it might be possible to establish a general conclusion and theory about the constructs in question and leadership success. Unless the conclusion is true with deductive certainty, it is considered an inductive conclusion.

The process of going from general to particular is also incorporated in the scientific method, especially in steps 3 to 6 (i.e., hypothesizing, predicting, and testing). A researcher with a conjectural statement could begin at step 3 of the scientific method and design empirical studies to test that conjecture and make predictions about the findings for particular datasets and studies.

It is also possible to start with a general corroborated scientific theory and then ask research questions and create hypotheses extending the theory to particular new settings. For example, a researcher might start with a theory that is established in the basic psychological research literature (e.g., social psychology) and see if it applies to a particular organizational issue. Alternatively, a researcher might begin with a theory that applies to a certain population and then test its validity in another population (e.g., examining whether the five-factor model of personality seen in Western samples applies to a newly studied applicant pool in a different country).

Overall, we believe that organizational researchers do not necessarily need to make the distinction between going to and from particular to general when conducting research. Instead, they can simply adopt the scientific method as it is used in other fields. Full adoption of the definition of the scientific method used in other fields would allow organizational researchers to make and publish research going from the particular to the general, without special calls for inductive research. It would also simplify the instruction of students. Learning the distinction between general and particular or inductive and deductive would no longer be needed. In addition, there would be greater consistency in how the scientific methods are taught in organizational sciences courses and courses in other areas of science.

The Scientific Method Incorporates the Philosophical Definitions of Induction and Deduction

In this section, we describe how the philosophical definitions of induction and deduction are incorporated in the scientific method. Induction deals with probabilities and it appears clearly between steps 5 and 6, during which a researcher makes a probabilistic determination (based on statistical significance and effect size) as to whether the hypothesis is supported. It also appears clearly in step 7, when a field of scientific researchers makes a determination as to whether there is a high enough probability that a hypothesis is true for it to be considered a theory. Bayesian statistics, whose creator was one of the principal contributors to the field of inductive logic (Fitelson, 2006), can play a key role in these determinations. Indeed, validity generalization (a scientific theory) uses Bayesian statistics (Schmidt & Hunter, 1977; Schmidt & Raju, 2007; Schmidt et al., 1979), and some authors have noted the ability of Bayesian statistics to synthesize prior knowledge, accept null hypotheses, and state theories (Jebb & Woo, 2015; Kruschke, Aguinis & Joo, 2012).

There are also some striking similarities between the outcomes of Bayesian statistical analyses (e.g., a probability) and the results of expectancy table analyses used in the personnel selection literature. For instance, given the meta-analytic operational validity of .66 for general mental ability tests in medium-complexity jobs, an individual in the top 25% of scores on this test has a 57.6% probability of being in the top 25% on the criterion (compared to 3.1% for individuals with a test score in the bottom 25%).[5] These types of probabilities are commonly used to predict risk in actuarial and medical settings, to create weather forecasts, and even predict earthquakes. However, in all instances, the conclusions involve a probability; even if the probability can be stated with extreme precision (e.g., the 95% confidence interval for the probability of 57.6% in the above example is 55.7% to 59.6%)[6], the conclusions about an individual case are still uncertain and involve a probability. Essentially, at a global level, all empirical research is inductive and none is exclusively deductive because the researcher is always making an inference based on incomplete information and the conclusion can only be probabilistic.

Although all empirical research in psychology is inductive, there are examples of truly deductive research in these fields. Mathematics is a field that makes extensive use of deduction, especially in mathematical proofs and theorems (which are essentially theories based on deduction). Consider the mathematical proof for the correction of a correlation coefficient between a predictor (*p*) and a criterion (*c*) for criterion unreliability. Using a deductive process, it is possible to go from the equations for a partial correlation coefficient and a variance decomposition to the equation for the correction for criterion unreliability. Assuming that the premises are true (e.g., the formulas for partial correlation and variance decomposition are correct, the algebraic manipulations used in the derivation are correct, measurement error is random and uncorrelated with the predictor and criterion), the conclusion (i.e., the formula for correction for unreliability is $r_{PC}/\sqrt{rel_C}$) must be true. This type of reasoning is prominent in the psychometric literature. In many ways, psychometricians could be called theoretical psychologists and the derivations and proofs for classical test theory, item-response theory, and generalizability theory are examples of deductive reasoning.

There are also instances in which either of the philosophical definitions of induction and deduction can appear in the scientific method. Consider step 1 of the scientific method (making an observation). Whether induction or deduction is involved depends on the nature of the observation itself. An observation that entails complete information on the population of interest could be the basis for a deductive conclusion. For instance, an organizational scientist might observe that all of the employees in an organization who were rated unsuccessful were trained at training center X. If the entire population was studied here, a valid deductive conclusion would be that if employee Y was rated unsuccessful, then employee Y was trained at training center X. An observation that would be the basis for an inductive conclusion would be one in which the organizational scientist observed that 80% of the employees rated unsuccessful were trained at training center X. A valid inductive inference would be that if employee Y was rated unsuccessful, then there is a .8 probability that employee Y was trained at training center X.

Conclusion

The scientific methodology of I-O psychology and management research is currently drifting away from the scientific method and toward an embrace of pseudotheory. Our research culture is impeding adoption of the scientific method. Researchers are motivated to make theoretical contributions in their papers, yet by doing so they have adopted an incorrect viewpoint of what theory is. According to the scientific method, making a theoretical contribution in a single primary study is impractical if not impossible. If the research literatures of I-O psychology and management are to truly make a contribution to scientific knowledge, major changes are needed. (Re)Adopting the scientific method provides a sound and established basis for further advancing I-O and management research and making better contributions to organizations. It also provides a framework for how data, methods, and theory are necessarily related in scientific research.

Authors' Note

The views expressed in this chapter are those of the authors and do not necessarily reflect the views of U.S. Customs and Border Protection or the U.S. Federal Government. The authors would like to thank Magda Colberg, Kevin Murphy, and Paul Sparrow for their valuable comments and suggestions on this chapter. Portions of this chapter are based on an invited Distinguished Early Career Contributions Award address given by the first author at the 31st meeting of the Society for Industrial and Organizational Psychology, Anaheim, CA.

Notes

1 The low level of statistical power in many psychological studies is also unfortunate. Thirty years ago, Cohen (1992) lamented that there was little increase in statistical power of psychology studies since the first edition of his power handbook (Cohen, 1969) was published. Low statistical power continues to be an issue in the literature as the Open Science Collaboration (2015) reported a median sample size of 54 cases for the 97 original psychological studies it replicated. The mean replication effect size was a correlation of .197. Thus, on average, the authors of the original studies were attempting to detect a correlation of .197 with only 54 cases, which equates to a power of .30 according to G*POWER (Faul et al., 2007; Faul et al., 2009).

2 Traditionally, a laboratory notebook consists of a bound book with blank pages in which a researcher documents their empirical research and analyses. Writing is often done by hand in ink and the pages cannot be removed inconspicuously. This allows for a diary or journal of the researcher's activities and can be made available to other researchers for inspection. Computer-based versions of laboratory notebooks are also used.

3 This might be due to the perceived complexity of human behavior. Oftentimes we hear that organizational behavior is the study of people (in organizations) and that an individual human's behavior is complex, thus requiring complicated psychological models. However, psychology is not all that different from the other sciences in regard to the magnitude of empirical findings, the precision of measurements, the consistency of results, and the accuracy in predicting individual-level outcomes (Hedges, 1987; Meyer et al., 2001). Indeed, very few sciences can predict individual-level outcomes with complete accuracy. Consider the difficulty a meteorologist faces predicting the weather on any given day with complete accuracy. Actuaries predicting whether or not an individual will file an insurance claim and credit bureaus predicting individuals' creditworthiness also have difficulty predicting future events with complete certainty. Even in the laboratory sciences, difficulties exist. Most chemical reactions do not result in all molecules/atoms reacting (i.e., the "actual yield" of a most chemical reactions is less than the "theoretical yield"). Although there are many causes of an individual's behavior, there are also many causes of the behavior of an individual molecule, the weather, the economy, and so forth. In other fields, it is often helpful to study the effects of an independent variable on a dependent variable using parsimonious models and theories. We suggest that the same process be used by organizational researchers. Attempting to create a model that explains everything using every possible independent variable leads to a model that is difficult to test and to validate. It also introduces the possibility of redundancy and overlap in the independent variables (e.g., does the second independent variable relate to the dependent variable because it is a proxy for the first independent variable or are both measures of the same construct?).

4 This incorrect definition persists in the psychometric literature and appears in Carroll's (1993) treatise and McGrew's (2009; McGrew & Evans, 2004) reviews. Although using the incorrect definition of induction, Carroll (p. 238) noted the work of Colberg and her colleagues on general-to-particular inductive tasks.

5 These values were obtained using Hunter et al.'s (2006) reanalysis of Hunter's (1980) meta-analysis analysis and syntax for computing expectancies from Cucina, Berger & Busciglio, (2017).

6 This is based on the expectancies (obtained using Cucina et al.'s, 2017, R syntax) for the 95% confidence interval about the point estimate of .66 (.63 to .69 using the sample size of 12,933 reported by Hunter, 1986, and applying the corrections for unreliability and range restriction to the uncorrected upper and lower bounds of the confidence interval).

References

Aguinis, H., Bradley, K. J., & Broderson, A. (2014). Industrial-organizational psychologists in business schools: Brain drain or eye opener? *Industrial and Organizational Psychology*, 7(3), 284–303.

Aguinis, H., Mazurkiewicz, M. D., & Heggestad, E. D. (2009). Using web-based frame-of-reference training to decrease biases in personality-based job analysis: An experimental field study. *Personnel Psychology*, *62*(2), 405–438.

Blunt, C. J. (2015). Hierarchies of evidence in evidence-based medicine (Unpublished doctoral dissertation). London, UK: London School of Economics.

Blunt, C. J. (2020, August 10). *Hiearchies of evidence*. London, UK: London School of Economics. Retrieved December 12, 2020, from http://cjblunt.com/hierarchies-evidence/

Bradley, J. C. (2007). Open notebook science using blogs and wikis. *Nature Precedings*. https://doi.org/10.1038/npre.2007.39.1

Bradley, J. C. (September 26, 2006). Open notebook science. *Drexel CoAS E-Learning*. Retrieved August 11, 2021, from http://drexel-coas-elearning.blogspot.com/2006/09/open-notebook-science.html

Brooker, R. J., Widmaier, E. P., Graham, L. E., & Stiling, P. D. (2020). *Biology*. (5th ed.). New York, NY: McGraw-Hill.

Brush, S. G. (1999). Why was relativity accepted? *Physics in Perspective*, *1*, 184–214.

Burgus, R., Dunn, T. F., Desiderio, D., Ward, D. N., Vale, W., & Guillemin, R. (1970). Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor. *Nature*, *226*(5243), 321–325.

Campbell, J. P., & Wilmot, M. P. (2018). The functioning of theory in industrial, work and organizational psychology (IWOP). *The SAGE handbook of industrial, work & organizational psychology, 3v: Personnel psychology and employee performance; organizational psychology; managerial psychology and organizational approaches*, 1.

Canadian Task Force on the Periodic Health Examination. (1979). The periodic health examination. *Canadian Medical Association Journal*, *121*, 1193–1254.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

Carter, N. T., Dalal, D. K., Boyce, A. S., O'Connell, M. S., Kung, M. C., & Delgado, K. M. (2014). Uncovering curvilinear relationships between conscientiousness and job performance: How theoretically appropriate measurement makes an empirical difference. *Journal of Applied Psychology*, *99*(4), 564–586.

Cartwright, N. (1999). *The dappled world: A study of the boundaries of science*. New York, NY: Cambridge University Press.

Chaisson, E., & McMillan, S. (2017). *Astronomy: A beginner's guide to the universe* (8th ed.). London, UK: Pearson Education, Inc.

Chang, R., & Goldsby, K. A. (2019). *Chemistry* (13th ed.). New York, NY: McGraw-Gill.

Cochrane Library. (n.d.). *About Cochrane reviews*. John Wiley & Sons, Inc. Retrieved December 22, 2020, from https://www.cochranelibrary.com/about/about-cochrane-reviews.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Colberg, M., Nester, M. A., & Trattner, M. H. (1985). Convergence of the inductive and deductive models in the measurement of reasoning abilities. *Journal of Applied Psychology*, *70*(4), 681.

Cook, T. D., & D. T. Campbell (1979). *Quasi experimentation: Design and analytical issues for field settings*. Chicago, IL: Rand McNally.

Coon, D., & Mitterer, J.O. (2018). *Psychology: Modules for active learning* (14th ed.). Stamford, CT: Cengage Learning.

Corbyn, Z. (2010, January 23). Publisher attempts to rein in radical medical journal: Editor rejects proposal to have submissions peer reviewed. *Times Higher Education*. Retrieved June 8, 2015, from https://www.timeshighereducation.co.uk/news/publisher-attempts-to-rein-in-radical-medical-journal/410113.article.

Corley, K. G., & Gioia, D. A. (2011). Building theory about theory building: What constitutes a theoretical contribution? *Academy of Management Review*, *36*(1), 12–32.

Cortina, J. M. (2016). Defining and operationalizing theory. *Journal of Organizational Behavior*, *37*(8), 1142–1149.

Cortina, J. M. (2014). In Z. Sheng (Interviewer). An interview with SIOP's newly elected President, Dr. Jose Cortina. *The I/ON: The official newsletter of the industrial/organizational psychology program at George Mason University*. Fairfax, VA: George Mason University. Retrieved February 7, 2014, from http://www.gmu.edu/org/iopsa/Fall2013ION_updatedFINALpdf.pdf.

Cortina, J. M., Aguinis, H., & DeShon, R. P. (2017). Twilight of dawn or evening? A century of research methods. *Journal of Applied Psychology*, *102*(3), 274–290.

Crowell, B. (2020). *Light and Matter*. Fullerton, CA: Fullerton College.

Cucina, J. M., Nicklin, J. M., Ashkanasy, N., Cortina, J. M., Mathieu, J. E., & McDaniel, M. A. (2015). *The role of theory in industrial/organizational psychology research and practice*. Debate presented at the 30th meeting of the Society for Industrial and Organizational Psychology, Philadelphia, PA.

Cucina, J. M., Berger, J. L., & Busciglio, H. H. (2017). Creating expectancy charts: A new approach. *Personnel Assessment and Decisions*, *3*(1), 1.

Cucina, J. M., & Hayes, T. L. (2015). Comment on estimating the reproducibility of psychological science. *Science*.

Cucina, J. M., & Moriarty, K. O. (2015). A historical look at theory in industrial-organizational psychology journals. *The Industrial-Organizational Psychologist*, *53*(1), 57–70.

Cucina, J. M., Hayes, T. L., Walmsley, P. T., & Martin, N. R. (2014). It is time to get medieval on the overproduction of pseudotheory: How Bacon (1267) and Alhazen (1021) can save I/O psychology. *Industrial and Organizational Psychology: Perspectives on Science and Practice,* 7(3), 356–364.

Cucina, J. M., Martin, N. R., Vasilopoulos, N. L., & Thibodeaux, H. F. (2012). Self-serving bias effects on job analysis ratings. *The Journal of Psychology: Interdisciplinary and Applied*, *146*(5), 1–21.

Cucina, J. M., Su, C., Busciglio, H. H., & Thompson Peyton, S. (2015). Something more than *g*: Meaningful Memory uniquely predicts training performance. *Intelligence*, *49*, 192–206.

Cucina, J. M., Vasilopoulos, N. L., & Sehgal, K. (2005). Personality-based job analysis and the self-serving bias. *Journal of Business and Psychology*, *20*(2), 275–290.

Cumpston, M., & Chandler J. (2020). Chapter IV: Updating a review. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, T. M. J. Page, & V. A. Welch (Eds.). *Cochrane handbook for systematic reviews of interventions (Version 6.1; updated September 2020)*. Hoboken, NJ: John Wiley & Sons & London, UK: Cochrane Collaboration. Available from www.training.cochrane.org/handbook.

Davis, M. S. (1971). That's interesting! Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy of the social sciences*, *1*(2), 309–344.

Dilchert, S. (2017). Future of research published in the International Journal of Selection and Assessment: Incoming editor's perspective. *International Journal of Selection and Assessment*, *25*(4), 416–418.

Dunning, D., Perie, M., & Story, A. L. (1991). Self-serving prototypes of social categories. *Journal of Personality and Social Psychology*, *61*(6), 957–968.

Edwards, J. R., Berry, J., & Kay, V. S. (2014, January). Bridging the great divide between theoretical and empirical management research. In *Academy of management proceedings* (Vol. 2014, No. 1, p. 17696). Academy of Management.

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, *37*(7), 1097–1126.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149–1160.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Fitelson, B. (2006). Inductive Logic. In J. Pfeifer & S. Sarkar (Eds.). *Philosophy of science: An encyclopedia*. London, UK: Routledge Press.

Freeman, S., Quillin, K., Allison, L., Black, M., Podgorski, G., Taylor, E., & Carmichael, J. (2017). *Biological science* (6th ed.). Boston, MA: Pearson Education, Inc.

Gilbert, T. R., Kirss, R.V., Foster, N., & Davies, G. (2014). *Chemistry: An atoms-focused approach*. New York, NY: W.W. Norton & Company, Inc.

Goldberg, S. (1987). Putting new wine in old bottles: The assimilation of relativity in America. (pp. 1–26). In T. F. Glick. (Ed.). *The comparative reception of relativity*. Boston, MA: D. Reidel Publishing Company.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: SAGE Publications, Inc.

Hambrick, D. C. (2007). The field of management's devotion to theory: Too much of a good thing? *Academy of Management Journal*, *50(6),* 1346–1352.

Harvey, A., & Schucking, E. (2000). Einstein's mistake and the cosmological constant. *American Journal of Physics*, *68*(8), 723–727.

Hawthorne, J. (2017). Inductive Logic. *The Stanford Encyclopedia of Philosophy*. (Spring 2017 ed.), In E.N. Zalta (Ed.). https://plato.stanford.edu/archives/spr2017/entries/logic-inductive

Hedges, L.V. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. *American Psychologist*, *42*(2), 443–455.

Hewitt, P. G. (2015). *Conceptual physics* (12th ed.). Boston, MA: Pearson.

Highhouse, S. (2014). Do we need all these words? The need for new publishing norms in I-O psychology. *The Industrial-Organizational Psychologist*, *51*(3), 83–84.

Hockenbury, S. E., Nolan, S., & Hockenbury, D. H. (2015). *Psychology* (7th ed.). New York, NY: Worth Publishers.

Hoefnagels, M. (2019). *Biology: The essentials* (3rd ed.). New York, NY: McGraw-Hill Education.

Howick, J., Glasziou, P., & Aronson, J. K. (2010). Evidence-based mechanistic reasoning. *Journal of the Royal Society of Medicine*, *103*, 433–441.

Hunter, J. E. (1980). *Validity generalization for 12,000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB)*. Washington, DC: U.S. Department of Labor, Employment Service.

Internet Encyclopedia of Philosophy. (2016). *Deductive and Inductive Arguments*. Retrieved July 17, 2016, from http://www.iep.utm.edu/ded-ind/.

Jebb, A. T., & Woo, S. E. (2015). A Bayesian primer for the organizational sciences: The "two sources" and an introduction to BugsXLA. *Organizational Research Methods*, *18*(1), 92–132.

Journal of Applied Psychology. (2004). Instructions to authors. [Editorial]. *Journal of Applied Psychology*, *89*(1), 178.

Kacmar, K. M., & Whitfield, J. M. (2000). An additional rating method for journal articles in the field of management. *Organizational Research Methods, 3*(4), 392–406.

Kay, L., Palen, S., Smith, B., & Blumenthal, G. (2013). *21st century astronomy* (4th ed.) New York, NY: W.W. Norton & Company, Inc.

Kaye, S. (n.d.). William of Ockham (Occam, c. 1280—c. 1349). *Internet Encyclopedia of Philosophy.* Retrieved January 4, 2021, from https://iep.utm.edu/ockham/#H2.

Kenny, D. A. (1979). *Correlation and causality*. New York: Wiley.

Kenny, D. A. (2004). *Correlation and causality* (Revised ed.). Storrs, CT: Author.

Kepes, S., & McDaniel, M. A. (2013). How trustworthy is the scientific literature in I-O psychology? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *6*, 252–268.

Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, *2*(3), 196–217.

Ketchen, D. J. (2002). Some candid thoughts on the publication process. *Journal of Management*, *28*, 585–590.

Klein, K. J., & Zedeck, S. (2004). Introduction to the special section on theoretical models and conceptual analyses – Theory in applied psychology: Lessons (re)learned. *Journal of Applied Psychology*, *89*(6), 931–933.

Köhler, T., DeSimone, J. A., & Schoen, J. L. (2020). Prestige does not equal quality: Lack of research quality in high-prestige journals. *Industrial and Organizational Psychology*, *13*, 321–327.

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods*, 15, 722–752.

LaCour, M. J., & Green, D. P. (2014). When contact changes minds: An experiment on transmission of support for gay equality. *Science*, *346*(6215), 1366–1369.

Lance, C. E., & Vandenberg, R. J. (2015). *More statistical and methodological myths and urban legends*. New York, NY: Routledge.

Licht, D., & Hull, M. (2014). *Scientific American: Psychology*. New York, NY: Worth Publishers.

Lilienfeld, S. O., Lynn, S. J., Ruscio, J., & Beyerstein, B. L. (2009). *50 great myths of popular psychology: Shattering widespread misconceptions about human behavior*. Malden, MA: Wiley-Blackwell.

Locke, E. A. (2007). The case for inductive theory building. *Journal of Management*, *33*(6), 867–890.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717.

Locke, K., & Golden-Biddle, K. (1997). Constructing opportunities for contribution: Structuring intertextual coherence and "problematizing" in organizational studies. *Academy of Management Journal*, *40*(5), 1023–1062.

Locke, E. A., Williams, K. J., & Masuda, A. (2015). The virtue of persistence. *The Industrial-Organizational Psychologist*, *52*(4), 104–5.

Lutgens, F. K., Tarbuck, E. J., & Tasa, D.G. (2014). *Essentials of Geology* (13th ed.). New York NY: Pearson.

Mason, K. A., Losos, J. B., Duncan, T. Welsh, C. J., Raven, P. H., & Johnson, G. B. (2020). *Biology.* (12th ed.). New York, NY: McGraw-Hill.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1–10.

McGrew, K. S., & Evans, J. J. (2004). *Internal and external factorial extensions to the Cattell-Horn-Carroll (CHC) theory of cognitive abilities: A review of factor analytic research since Carroll's seminal 1993 treatise.* (Carroll Human Cognitive Abilities Project Research Report No. 2). Institute for Applied Psychometrics.

Medical Hypotheses. (n.d.). Guide for authors. Retrieved December 22, 2020, from https://www.elsevier.com/journals/medical-hypotheses/0306-9877/guide-for-authors.

Merry, S. N., Hetrick, S. E., Cox, G. R., Brudevold-Iversen, T., Bir, J. J., & McDowell, H. (2012). Cochrane Review: Psychological and educational interventions for preventing depression in children and adolescents. *Evidence-Based Child Health: A Cochrane Review Journal*, 7(5), 1409–1685.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., & Reed, G. M. (2001). Psychological testing and psychological assessment: A review of evidence and issues. *American psychologist*, *56*(2), 128–165.

Mintzberg, H. (2005). Developing theory about the development of theory. *Great minds in management: The process of theory development*, 355–372.

Mischel, W. (1968). *Personality and Assessment*. New York, NY: John Wiley & Sons, Inc.

Monroe, J. S., & Wicander, R. (2015). *The changing earth: Exploring geology and evolution* (7th ed.). Stamford, CT: Cengage Learning.

Mupepele, A. C., Walsh, J. C., Sutherland, W. J., & Dormann, C. F. (2016). An evidence assessment tool for ecosystem services and conservation studies. *Ecological Applications*, *26(*5), 1295–1301.

National Academy of Sciences and Institute of Medicine. (2008). *Science, evolution, and creationism*. Washington, DC: The National Academies Press.

National Academy of Sciences. (n.d.). *Evolution resources at the National Academies: Definitions of evolutionary terms*. Washington, DC: Author. Retrieved March 6, 2021, from https://www.nationalacademies.org/evolution/definitions.

National Center for Science Education. (2016). *Definitions of fact, theory, and law in scientific work*. Oakland, CA: Author. Retrieved March 6, 2021 from https://ncse.com/library-resource/definitions-fact-theory-law-scientific-work.

Nature. (2015). For authors: Other material published in *Nature*. Retrieved September 21, 2015, from http://www.nature.com/nature/authors/gta/others.html.

Nicklin, J. M., & Spector, P. E. (2016). Point/Counterpoint introduction: The future of theory in organizational behavior research. *Journal of Organizational Behavior*, *37*(8), 1113–1115.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.

O'Boyle Jr., E. H., Banks, G. C., & Gonzalez-Mulé, E. (2017). The chrysalis effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management*, *43*(2), 376–399.

Olenick, J., Walker, R., Bradburn, J., & DeShon, R. P. (2018). A systems view of the scientist–practitioner gap. *Industrial and Organizational Psychology*, *11*(2), 220–227.

Ones, D. S., Kaiser, R. B., Chamorro-Premuzic, T., & Svensson, C. (2017). Has industrial-organizational psychology lost its way? *The Industrial-Organizational Psychologist*, *54*(4). Retrieved from http://www.siop.org/tip/april17/lostio.aspx.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716.

Pain, E. (2019, September 3). How to keep a lab notebook. *Science*. Retrieved August 11, 2021, from https://www.sciencemag.org/careers/2019/09/how-keep-lab-notebook.

Pauling, L., & Corey, R. B. (1953). Structure of the nucleic acids. *Nature*, *171*(4347), 346.

Pfeffer, J., & Sutton, R. I. (2006). Evidence-based management. *Harvard Business Review*, *84*(1), 62–74.

Phelan, J. (2015). *What is life? A guide to biology* (3rd ed.). New York, NY: W.H. Freeman and Company.

Platt, J. R. (1964). Strong inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. *Science*, *146*, 347–353.

Popper, K. (1934). *Logik der forschung: Zur erkenntnistheorie der modernen naturwissenschaft*. Vienna, Austria: Verlag von Julius Springer.

Popper, K. (1959). *The logic of scientific discovery*. London, UK: Hutchinson and Company.

Reay, T., Berta, W., & Kohn, M. K. (2009). What's the evidence on evidence-based management? *Academy of Management Perspectives*, *23*(4), 5–18.

Reay, T., Zafar, A., Monteiro, P. & Glaser, V. (2019). Presenting findings from qualitative research: One size does not fit all!. In T. Zilber, J. Amis and J. Mair. *The Production of Managerial Knowledge and Organizational Theory: New Approaches to Writing, Producing and Consuming Theory*. Emerald Publishing Limited.

Roberts, J. S., Donoghue, J. R., & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, *24*, 3–32.

Rogers, K. (2016, November 22). Scientific hypothesis. *Encyclopaedia Britannica*. Chicago, IL: Encyclopaedia Britannica, Inc. Retrieved January 24, 2017, from: https://www.britannica.com/topic/scientific-hypothesis.

Ryan, A. M., & Ployhart, R. E. (2014). A century of selection. *Annual Review of Psychology*, *65*, 693–717.

Saylors, R., & Trafimow, D. (2021). Why the increasing use of complex causal models is a problem: On the danger sophisticated theoretical narratives pose to truth. *Organizational Research Methods*, 24(3), 616–629.

Schapira, M., The Open Lab Notebook Consortium, Harding, R., J. (2019). Open laboratory notebooks: Good for science, good for society, good for scientists. [Version 2; peer review: 2 approved, 1 approved with reservations]. *F1000Research*, *8*, 87. pmid:31448096.

Schmidt, F. L. (2015). History and development of the Schmidt–Hunter meta-analysis methods. *Research synthesis methods*, *6*(3), 232–239.

Schmidt, F. L., & Hunter, J. E. (1977). Development of a general solution to the problem validity generalization. *Journal of Applied Psychology*, *62*, 529–540.

Schmidt, F. L., & Hunter, J. E. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173.

Schmidt, F. L., & Raju, N. S. (2007). Updating meta-analytic research findings: Bayesian approaches versus the medical model. *Journal of Applied Psychology*, *92*(2), 297.

Schmidt, F. L., Hunter, J. E., Pearlman, K., & Shane, G. S. (1979). Further tests of the Schmidt-Hunter Bayesian validity generalization model. *Personnel Psychology*, *32*, 257–281.

Schmidt, F., & Hunter, J. (2003). History, development, evolution, and impact of validity generalization and meta-analysis methods, 1975–2001. In K.R. Murphy (Ed.). *Validity generalization: A critical review*. (pp. 42–76). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schmidt, F.L., Hunter, J. E., Pearlman, K., & Hirsh, H. R. (1985). Forty questions about validity generalization and meta-analysis. *Personnel Psychology*, *38*, 697–798.

Schmidt, F. L., Hunter, J. E., & Urry, V.W. (1976). Statistical power in criterion-related validation studies. *Journal of Applied Psychology*, *61*(4), 473–485.

Schneider, B. (2018). Being competitive in the talent management space. *Industrial and Organizational Psychology*, *11*(2), 231–236.

Shen, W., Cucina, J. M., Walmsley, P., & Seltzer, B. (2014). When correcting for unreliability of job performance ratings, the best estimate is still .52. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 7, 519–524.

Shepherd, D. A., & Suddaby, R. (2017). Theory building: A review and integration. *Journal of Management*, *43*(1), 59–86.

Shojania, K. G., Sampson, M., Ansari, M. T., Ji, J., Doucette, S., & Moher, D. (2007). How quickly do systematic reviews go out of date? A survival analysis. *Annals of Internal Medicine*, *147*(4), 224–233.

Siegel, E. (2013, May 17). "Einstein's greatest blunder" was REALLY a blunder! *ScienceBlogs.* Retrieved December 23, 2020, from https://scienceblogs.com/startswithabang/2013/05/17/einsteins-greatest-blunder-was-really-a-blunder.

Silberberg, M., & Amateis, P. (2018). *Chemistry: The molecular nature of matter and change* (8th ed.). New York, NY: McGraw-Hill Education.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Skyrms, B. (2000). *Choice and chance: An introduction to inductive logic* (4th ed.). Wadsworth, Inc.

Spector, P. E., Rogelberg, S. G., Ryan, A. M., Schmitt, N., & Zedeck, S. (2014). Moving the pendulum back to the middle: Reflections on and introduction to the inductive research special issue. *Journal of Business and Psychology*, *29*(4), 499–502.

Stanovich, K. E. (2010). *How to think straight about psychology* (9th ed.). Boston, MA: Allyn & Bacon

Sutton, R. I., & Staw, B. M. (1995). What theory is *not*. *Administrative Science Quarterly*, *40*, 371–384.

Texas Gateway/Texas Education Agency. (n.d.). Theories. *Texas Gateway for online resources*. Austin, TX: Texas Education Agency. Retrieved March 6, 2021, from https://www.texasgateway.org/resource/theories.

Thurstone, L. L. (1938). Primary mental abilities. *Psychometric Monographs*, (No. 1).

Trafton, L. M., & Wildey, R. L. (1970). Jupiter: His limb darkening and the magnitude of his internal energy source. *Science*, *168*(3936), 1214–1215.

U. S. Department of Labor. (1958). *Guide to the use of General Aptitude Test Battery – Section III: Development*. Washington, DC.

Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *20*(1), 57–59.

Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, *81*(5), 557–574.

Watson, J. D. (1997). *The double helix*. *London*: *Weidenfeld & Nicholson*.

Watson, J. D., & Crick, F. H. (1953). Molecular structure of nucleic acids. *Nature*, *171*(4356), 737–738.

Wicander, R., & Monroe, J. S. (2013). *GEOL2*. Belmont, CA: Brooks/Cole.

Woo, S. E., O'Boyle, E. H., & Spector, P. E. (2017). Best practices in developing, conducting, and evaluating inductive research. *Human Resources Management Review*, *27*, 255–264.